08 474450 A/NO FEE

0372/0B128

UNIVERSAL PRIMER SEQUENCE FOR MULTIPLEX DNA AMPLIFICATION

Field of the Invention

This invention pertains to universal primers having use in amplification of DNA sequences by methods such as polymerase chain reaction (PCR), specifically to primers that allow the simultaneous amplification of a multiplicity of DNA sequences.

Background of the Invention

Polymerase chain reaction (PCR) is a method whereby virtually any DNA sequence can be selectively amplified. The method involves using paired sets of oligonucleotides of predetermined sequence that hybridize to opposite strands of DNA and define the limits of the sequence to be amplified. The oligonucleotides prime multiple sequential rounds of DNA synthesis catalyzed by a thermostable DNA polymerase. Each round of synthesis is typically separated by a melting and re-annealing step, allowing a given DNA sequence to be amplified several hundred-fold in less than an hour (Saiki et al., *Science* **239**:487, 1988).

The simplicity and reproducibility of these reactions has given PCR broad applicability. For example, PCR has gained widespread use for the diagnosis of inherited

disorders and susceptibility to disease. Typically, the genomic region of interest is amplified from either genomic DNA or from a source of specific cDNA encoding the cognate gene product. Mutations or polymorphisms are then identified by subjecting the amplified DNA to analytical techniques such as DNA sequencing, hybridization with allele specific oligonucleotides, restriction endonuclease cleavage or single-strand conformational polymorphism (SSCP) analysis.

For the analysis of small genes or genes where the mutant allele or polymorphism is well characterized, amplification of single defined regions of DNA is sometimes sufficient. When analyzing large and/or undefined genes, however, multiple individual PCR reactions are often required to identify critical base changes or deletions. Thus, to streamline the analysis of large complex genes, multiplex PCR (i.e., the simultaneous amplification of different target DNA sequences in a single PCR reaction) has been utilized.

The results obtained with multiplex PCR are, however, frequently complicated by artifacts of the amplification procedure. These include "false-negative" results due to reaction failure and "false-positive" results such as the amplification of spurious products, which may be caused by annealing of the primers to sequences which are related to, but distinct from, the true recognition sequences.

For use in multiplex PCR, a primer should be designed so that its predicted hybridization kinetics are similar to those of the other primers used in the same multiplex reaction. While the annealing temperatures and primer concentrations may be calculated to some degree, conditions generally have to be empirically determined for each multiplex reaction. Since the possibility of non-specific priming increases with each additional primer pair,

3

conditions must be modified as necessary as individual primer sets are added. Moreover, artifacts that result from competition for resources (e.g., depletion of primers) are augmented in multiplex PCR, since differences in the yields of unequally amplified fragments are enhanced with each cycle. Given these limitations, the development of a new diagnostic test can be very labor- intensive and costly.

Weighardt et al. (*PCR Methods and App.* **3**:77, 1993) describe the use of 5'-tailed oligonucleotides for PCR. However, a key feature of this amplification method involves separate annealing and primer extension reactions for each primer, which is not practical in a multiplex context.

Thus, there is a need in the art for primers that allow multiplex PCR reactions to be designed and carried out without elaborate optimization steps, irrespective of the potentially divergent properties of the different primers used. Furthermore, there is a need in the art for primers that allow multiplex PCR reactions that simultaneously produce equivalent amounts of each one of many amplification products.

Summary of the Invention

This invention pertains to primers that allow simultaneous amplification of multiple DNA target sequences present in a DNA sample. According to the invention, the DNA sample in a single reaction mixture is contacted with a multiplicity of paired oligonucleotide primers having the structure 5'-XY-3', wherein: X comprises a sequence that does not hybridize to the target sequence; the melting temperature of a hybrid between X and its complement in the

absence of other sequences is greater than about 60°C; and Y comprises a sequence contained within or flanking the target sequence or its complement.

Multiple cycles of melting, reannealing, and DNA synthesis (i.e., a PCR reaction) are thereafter performed with the above mentioned DNA sample and the oligonucleotide primers. Preferably, X comprises the sequence 5'-GCGGTCCCAAAAGGGTCAGT-3'. Amplified target sequences may then be detected by any method, including, for example, hybridization with allele-specific oligonucleotides, restriction endonuclease cleavage, or single-strand conformational polymorphism (SSCP) analysis.

The invention also encompasses a method for detecting multiple defined target DNA sequences in a DNA sample. This method is carried out by performing the same procedure set forth above, in which the 3' sequence of one primer of each pair comprises a target DNA sequence itself or its complement. The method includes a further step of detecting the amplification products, preferably by gel electrophoresis. In this embodiment, the presence or absence of an amplification product is diagnostic of the presence or absence of the target sequence in the original DNA sample.

In another aspect, the invention encompasses methods for high-throughput genetic screening. The method, which allows the rapid and simultaneous detection of multiple defined target DNA sequences in DNA samples obtained from a multiplicity of individuals, is carried out by simultaneously amplifying many different target sequences from a large number of patient DNA samples, using oligonucleotide primer pairs as above.

In yet another aspect, the present invention provides single-stranded oligonucleotide DNA primers for amplification of a target DNA sequence in a multiplex

polymerase chain reaction. The primers have the structure 5'-XY-3', wherein X comprises the sequence 5'-GCGGTCCCAAAAGGGTCAGT-3' (SEQ ID NO: 65), and Y comprises a sequence contained within or flanking a target sequence or its complement. Typically, Y comprises a sequence from 17 to 25 bases in length, and the melting temperature of hybrids between the primers and their complements is at least 72°C in 0.5M NaCl.

The methods and compositions of the present invention can be applied to the diagnosis of genetic and infectious diseases, gender determination, genetic linkage analysis, and forensic studies.

Brief Description of the Drawings

Figures 1A and 1B are (SEQ ID NO: 1 through SEQ ID NO: 64) ~~Figure 1 is~~ a table listing amplicon-specific oligonucleotide primer sequences.

Figures 2A-2D are ~~Figure 2~~ is an illustration of an agarose gel in which PCR amplification products corresponding to exon 21 of the CFTR gene are resolved.

Figure 3 is an illustration of an agarose gel in which PCR amplification products corresponding to exons 3, 4, 5, 7, 9, 10, 11, 12, 13, 14b, 17b, 19, 20, 21, and intron 19 of the CFTR gene are resolved. Lanes 1-8 represent products derived from genomic DNA samples isolated from blood cells, while in lanes 9-12 the genomic DNA template was derived from buccal cells. Lanes 1-4 and 9-12 show the amplification products obtained using chimeric primers according to the present invention; lanes 5-8 and 13-16 show amplification products obtained using sequence-specific primers. Lane M shows the electrophoretic pattern of ϕX174 Hae III-digested marker DNA. The correspondence between exon designation and size is shown Figures 1A and 1B ~~in Figure 1.~~

Figure 4 is an illustration of an agarose gel in which the products of a single PCR amplification assay for multiple gene loci are resolved. The loci correspond to the CFTR locus as in Figure 3 (lanes 1-6); α-galactosidase and sickle-cell genes (lanes 7-12); α-galactosidase and Tay-Sachs genes (lanes 13-18), and β-thalassemia (lanes 19-24). In lanes 1-3, 5-8, 13-15 and 19-21, chimeric primers according to the present invention were used for amplification. In lanes 4-6, 9-12, 16-18 and 22-24, the corresponding sequence-specific non-chimeric primers were used.

Figure 5 is an illustration of an agarose gel in which PCR amplification products are resolved corresponding to different segments of the human WT-1 gene. The products shown in lanes 1, 2, 5, 6, 9, 10, 13, 14, 17, 18, 21, 22 were amplified with chimeric primer pairs designed according to the present invention. The products shown in lanes 3, 4, 7, 8, 11, 12, 15, 16, 19, 20, 23, 24 were amplified with the corresponding sequence-specific primers. Amplifications represent each of six amplicons within the WT-1 gene (B, F, H, J, N, and O; see Table 1). Lane M is ϕX174 Hae III-digested marker DNA.

Detailed Description of the Invention

All patent applications, patents, and literature references cited in this specification are hereby incorporated by reference in their entirety. In case of conflict, the present description, including definitions, will control.

Definitions:

1. "Amplification" of DNA as used herein denotes the use of polymerase chain reaction (PCR) to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. An "amplicon" is a target DNA sequence that is amplified by PCR.

2. "Multiplex PCR" as used herein refers to the simultaneous amplification of multiple DNA targets in a single polymerase chain reaction (PCR) mixture.

3. "High-throughput" denotes the ability to simultaneously process and screen a large number of DNA samples (e.g. in excess of 100 genomic DNAs) in a rapid and economical manner, as well as to simultaneously screen large numbers of different genetic loci within a single DNA sample.

The present invention encompasses methods and compositions that allow the efficient and essentially simultaneous amplification of different target DNA sequences in a single polymerase chain reaction (i.e., multiplex PCR). Preferably, equivalent amounts of each amplification product are obtained. The method utilizes novel chimeric oligonucleotide primers that circumvent the technical difficulties associated with multiplex PCR that result in unequal amplification of different target sequences in the same reaction mix.

For example, in a standard PCR reaction employing more than a single pair of oligonucleotide primers, the obligatory imposition of a single set of reaction conditions generally means that one of the primer sets will function more efficiently in priming, causing the target sequence specified by that set of primers to be selectively amplified in the early cycles of amplification. Furthermore, the more efficient primers will also be depleted from the reaction sooner than the less efficient ones, resulting in the increased accumulation of non-specific

amplification products in later cycles of amplification. Clearly, these problems are magnified when it is desired to use multiple primer pairs (>3-4) in a single reaction.

The methods and compositions of the present invention circumvent these problems by imposing a uniformly high degree of specificity on the annealing reactions that occur between different primers present in the mixture and their cognate target sequences in the DNA template. During the early cycles of amplification, products are synthesized that contain the chimeric primers on either end. The chimeric primers then serve as high stringency recognition sequences for subsequent rounds of amplification. This results in normalizing the annealing efficiency of different primers and their cognate target sequences, and thus also normalizes the degree of amplification of different targets.

Primer Design

Multiplex PCR according to present invention utilizes chimeric oligonucleotide primers that include two domains. The 5' "half" of each primer may comprise any sequence between 17 and 25 bases in length that is unrelated to the target DNA, and has the property of forming hybrids with relatively high melting temperatures (e.g., T_ms > 60°C in the absence of other sequences). In some applications, when the target DNA sequence is embedded in a sequence of low complexity (i.e., $<10^8$ bp), primers may be used that form hybrids with lower melting temperatures. In a preferred embodiment, the 5' sequence comprises 5'-GCGGTCCCAAAAGGGTCAGT-3' (SEQ ID NO:65). This sequence, which is designated as a "universal primer sequence" (UPS), is derived from the bacteriophage vector M13mp18 (Messing J., *Meth. Enzymol.* **101**:20, 1983).

The 3' "half" of each primer comprises a target-specific sequence, i.e., a sequence that is either present or potentially present in the target DNA or its complement. These 3' sequences may comprise without limitation any such sequence of 17-25 bases, and preferably 20 bases, irrespective of the melting temperatures of hybrids formed between the isolated sequence and its complement.

In one embodiment, the 3' half of the primer is intended to hybridize with a genomic sequence flanking the target sequence of interest; in this case, the primer is used to amplify the target sequence for subsequent diagnostic tests such as, e.g., hybridization with allele-specific oligonucleotides, restriction endonuclease cleavage, or single-strand conformational polymorphism (SSCP) analysis. For this purpose, the 3' half of the primer must correspond to a sequence *known* to be present in all DNA samples to be tested (or its complement). Non-limiting examples of 3' primer halves useful in practicing the present invention are shown in Figures 1A and 1B ~~Figure 1~~.

In another embodiment, the amplification reaction itself serves as the critical diagnostic step. In this case, the 3' sequence of the primer corresponds to a defined wild-type version of a particular amplicon or its complement (or to a variant version or its complement) whose presence or absence is being tested. When such allele-specific sequences are incorporated into chimeric PCR primers according to the present invention, and the chimeric primers are used in amplification reactions, the absence of a given amplification product is considered definitive for the absence of the allele in the DNA sample being tested.

For use in a given multiplex PCR reaction, target-specific primer sequences are typically analyzed as a group to evaluate the potential for fortuitous dimer formation between

10

different primers. This evaluation may be achieved using commercially available computer programs for sequence analysis, such as *Gene Runner*, Hastings Software Inc. Other variables, such as the preferred concentrations of Mg^{+2}, dNTPs, polymerase, and primers, are optimized using methods well-known in the art (Edwards et al., *PCR Methods and Applications* 3:565,1994).

DNA templates

Any DNA sample may be used in practicing the present invention, including without limitation eukaryotic, prokaryotic and viral DNA. In a preferred embodiment, the target DNA represents a sample of genomic DNA isolated from a patient. This DNA may be obtained from any cell source or body fluid. Non-limiting examples of cell sources available in clinical practice include blood cells, buccal cells, cervicovaginal cells, epithelial cells from urine, fetal cells, or any cells present in tissue obtained by biopsy. Body fluids include blood, urine, cerebrospinal fluid, semen and tissue exudates at the site of infection or inflammation. DNA is extracted from the cell source or body fluid using any of the numerous methods that are standard in the art. It will be understood that the particular method used to extract DNA will depend on the nature of the source. The preferred amount of DNA to be extracted for use in the present invention is at least 5 pg (corresponding to about 1 cell equivalent of a genome size of 4×10^9 base pairs).

Multiplex PCR reaction conditions

In practicing the present invention, a DNA sample is contacted with pairs of chimeric oligonucleotide primers under conditions suitable for polymerase chain reaction. Standard PCR reaction conditions may be used, e.g., 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 200 μM deoxynucleotide triphosphates (dNTPs), and 25-100 U/ml Taq polymerase (Perkin-Elmer, Norwalk, CT).

The concentration of each chimeric primer in the reaction mixture can range from about 0.05 to about 4 μM. The optimal concentration for primer is evaluated by performing single PCR reactions using each primer pair individually. Similarly, each primer pair is evaluated independently to confirm that all primer pairs to be included in a single multiplex PCR reaction require the same amplification conditions (i.e., temperature, duration of annealing and extension steps). It was found (see example below) that all chimeric primers containing the M13 derived UPS as the 5' half of their sequence could be used at a broad range of annealing temperatures (i.e., 50-60°C).

Multiplex PCR reactions are carried out using manual or automatic thermal cycling. Any commercially available thermal cycler may be used, such as, e.g., Perkin-Elmer 9600 cycler.

Finally, the reaction products are analyzed using any of several methods that are well-known in the art. Preferably, agarose gel electrophoresis is used to rapidly resolve and identify each of the amplified sequences. In a multiplex reaction, different amplified sequences are preferably of distinct sizes and thus can be resolved in a single gel. In one embodiment, the reaction mixture is treated with one or more restriction endonucleases prior to electrophoresis.

Alternative methods of product analysis include without limitation dot-blot hybridization with allele-specific oligonucleotides and SSCP.

The following examples are intended to further illustrate the present invention without limiting the invention thereof.

Example 1: Effect of Chimeric Primers on Efficiency and Specificity of Amplification

The following experiment was done to evaluate the effects on amplification of incorporating the M13 UPS sequence into PCR primers.

A. METHODS

Primer design:

Three sequence-specific primer pairs used to amplify Exon 21 of the cystic fibrosis transmembrane regulator (CFTR) gene (Kerem et al., *Science* **245**:1073, 1989) are shown in Figures 1A and 1B. A chimeric version of one of the three primers was synthesized containing the M13 UPS sequence 5'-GCGGTCCCAAAAGGGTCAGT-3' (SEQ ID NO: 65) immediately 5' to the illustrated sequences. The oligonucleotides were synthesized using conventional chemistry and were purified by high-performance liquid chromatography prior to use.

DNA preparation:

Whole blood samples were collected in high glucose ACD Vacutainers™ (Beckton Dickenson & Co., Franklin Lanes, NJ). Following centrifugation, the buffy coat was collected and lysed with two washes of a 10:1 (v/v) solution of 14 mM NH_2Cl and 1 mM $NaHCO_3$. The nuclei were harvested by centrifugation, resuspended in lysis buffer (10 mM Tris-HCl, pH 8.0, 0.4 M NaCl, 2 mM EDTA, 0.5% SDS, 500 μg/ml proteinase K) and incubated overnight at

37°C. Samples were then extracted with 1/4th volume of saturated NaCl, and the DNA was collected by ethanol precipitation. The final DNA pellet was washed with 70% ethanol, air dried and dissolved in TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA).

Amplification reactions:

For single amplifications, 50 μl reaction mixtures were prepared containing 2 μg of genomic DNA prepared as described above in 1X PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl), 200 μM dNTPs, and 2.5 units *Taq* polymerase (Perkin-Elmer, Norwalk, CT). Multiplex PCR reactions were carried in a volume of 100 μl under the same conditions, except that 10 units of *Taq* polymerase per reaction was used. Primer concentrations ranged from 0.25 to 1.0 μM. Amplifications were carried out using a Perkin-Elmer 9600 themocycler (Perkin-Elmer, Norwalk, CT) for 28 cycles with ramping (melting at 94°C for 10 s, annealing at 50°C, 55°C, 60°C, or 65°C for 10 s, and extension at 72°C for 10s).

After completion of the reaction, 8 μl of the reaction products were loaded directly onto a 2% ethidium bromide-stained agarose gel and subjected to electrophoresis at 250 volts for 90 minutes. The amplification products were visualized with a UV transilluminator and photographed with an Alpha Innotech IS-500 Digital Imaging System version 1.97 (Sun BIO Science, Inc., Branford, CT).

B. RESULTS

The efficiencies with which the three CFTR primer pairs (designated SS#1, SS#2, and SS#3) primed amplification varied with primer concentration and temperature of annealing (figures 2A-2D). ~~(Figure 2)~~. The primer concentrations were as follows: Lanes 1, 4, 7, and 10, 1.0 μM; lanes

2, 5, 8, and 11, 0.5 μM; and lanes 3, 6, 9, and 12, 0.25 μM. The temperatures of annealing were 50°C, 55°C, 60°C and 65°C, as indicated.

The SS#1 and SS#3 primers, for example, were noticeably inefficient at annealing temperatures above 60°C. The primer pair designated SS#3-UPS, which corresponds to the SS#3 primers having the M13 UPS sequence on their 5' termini, was highly efficient in priming at all temperatures tested; furthermore, few spurious amplification products were detected in reactions containing SS#3-UPS primers. By contrast, SS#2 primers gave spurious amplification products at all three temperatures below 65°C.

Example 2: **Comparison of Multiplex PCR Reactions Using CFTR Primer Pairs Lacking and Containing M13 UPS**

Fifteen primer pairs that were used to amplify sequences contained in different exons of the *CFTR* gene are shown in ~~Figure 1.~~ Figures 1A and 1B. A parallel set of primers was synthesized in which the M13 UPS sequence was present 5' to the *CFTR*-specific sequences.

DNA prepared as described in Example 1 was incubated simultaneously with all fifteen UPS-containing or -lacking primer pairs, and amplification reactions were carried out using identical conditions (e.g., 60°C annealing temperature). Identical reaction conditions, cycling times and primer concentrations were used for both primer sets.

As shown in Figure 3, PCR reactions using the standard sequence-specific primer pairs fail to generate a clear multiplex PCR profile of the *CFTR* locus. Specifically, several of the expected bands are clearly under-represented due to differential amplification (Figure 3, lanes 5-8 and lanes 13-16). In contrast, a clear multiplex profile is obtained when the *CFTR* locus is amplified with the corresponding UPS tagged primer pairs. The expected bands are clearly

15

prominent and the profiles are virtually free of contaminating non-specific products (Figure 3, lanes 1-4 and 9-12). Moreover, equivalent banding patterns are observed over an 8-fold range of template concentrations when the UPS tagged primer pairs are employed. Conversely, the amplification profile generated using the non-tagged standard primer sets is sensitive to variations in the template concentration as evidenced by the changes in the intensity of individual bands (Figure 3, lanes 5-8 and 13-16).

Example 3: Use of Multiplex PCR to Simultaneously Amplify Different Disease-related Sequences Under Identical Conditions

DNA was isolated as described in Example 1 above and subjected to multiplex PCR amplification using different combinations of UPS-tagged and untagged primers (Figures 1A and (B). ~~(Figure 1).~~

The banding patterns for the following primer sets are displayed in Figure 4: lanes 1-3 and 4-6, *CFTR* locus; lanes 7-9 and 10-12, α-galactosidase (Gaucher's disease, GCR, Kornreich et al., *Nucleic Acids Res.* **17**:3301, 1989) and Sickle Cell Anemia, (SCA, Navon et al., *Science* **243**:1471, 1989); lanes 13-15 and 16-18, GCR and Tay-Sachs (TS, Tanaka et al., *Am. J. Hum. Genet.* **46**:329, 1990); lanes 19-21 and 22-24, β-thalassemia. Amplification of the human *WT1* gene (Wilms tumor, Varanasi et al., *Proc.Natl.Acad.Sci.USA* **91**:3554, 1994) using 6 primer pairs is presented in Figure 5. For the multiplex PCR reactions displayed in Figure 4, the UPS tagged primer pairs generate only the desired bands. Figure 4 further demonstrates that the UPS tagged primers yield co-amplified products that are more uniform with respect to the band intensities than the corresponding products generated from the non-tagged sequence specific primers (lanes 1-6, 7-12 and 13-18).

16

For 13 of the 14 UPS-tagged primer pairs, the expected bands are clearly prominent and virtually free of spurious amplification products. With the exception of one primer pair (Figure 5, lanes 5 and 6), which does not generate a detectable product when the chimeric primer is employed, the presence of the UPS sequence enhances the yield of the respective PCR products (Figure 5, lanes 1-4, 9-12, 17-20 and 21-24).

17